旅游高峰期交通流量的多时间尺度预测方法及应用 *

李凌雁，赵　欣+，翁钢民

(燕山大学 经济管理学院，河北 秦皇岛 066004)

摘　要：高峰期的旅游交通会引发交通需求量的激增、改变正常交通的时空分布规律、影响出行者的出行行为。对旅游交通需求进行准确的短时预测难度很大，其时空分布规律更加难以估计。为了解决这一问题，提高预测精度，首先分析了旅游交通流的多时间尺度可预测性，进而运用多时间尺度预测方法对旅游交通流量进行了预测。结果表明，模型预测性能良好、精度较高，从分类和分层的角度构建旅游交通多时间尺度预测模型、分析旅游交通流量分布和旅游交通流量动态分配，能够为旅游高峰期的道路交通运行态势快速评估和制定应急交通组织管理方案提供精确参考。

关键词：旅游高峰期；交通流量预测；可预测性；多时间尺度

中图分类号：U495；TP391　　doi: 10.3969/j.issn.1001-3695.2018.09.0633

Multi-time scale prediction method and application on tourism fastigium traffic flow

Li Lingyan, Zhao Xin+, Weng Gangmin

(*School of Economics & Management, Yanshan University, Qinhuangdao* Hebei 066004, *China*)

Abstract: Peak tourist traffic will lead to sharp increasing of traffic demand, change the temporal and spatial distribution of normal traffic, and influence traffic behavior. It is very difficult to predict the short-time travel traffic demand and its temporal and spatial distribution. In order to solve this problem and improve the prediction accuracy, this paper analyzes the predictability of multi-time scale of tourism traffic flow, and then the multi-time scale forecasting method is used to predict tourism traffic flow. The results show that the model has good prediction performance and high accuracy. From the perspective of classification and stratification, the multi-time scale prediction model of tourism traffic was constructed, and the distribution of tourism traffic flow and dynamic distribution of tourism traffic flow were analyzed, which provided precision reference for the rapid assessment of road traffic operation situation in the peak travel season and the formulation of emergency traffic organization and management.

Key words: tourism fastigium; traffic flow forecasting; predictability; multi-time scale

0　引言

旅游交通系统是一个复杂系统，在旅游城市现有旅游交通硬件设施的基础上，如何通过应用技术的升级保证旅客顺利、舒适、高效完成旅游活动，是当前旅游交通科学领域持续研究的热点问题[1,2]。目前，我国城市交通问题日益严重，交通拥堵、事故频发、尾气污染等问题一定程度上制约了城市旅游的可持续发展[3]。在旅游交通环境下，城市交通的常态化问题被进一步放大，严重影响了游客的正常旅游需求，旅游效率低下，严重时会引发影响社会安定的群体事件[4]。因此准确掌握和预测旅游交通流量对于提前做好交通组织与管理、保证交通顺畅、提高旅游交通服务质量、保证旅游者的时间效益最大化具有重要意义。

旅游交通管理中多模式最优路径的优化、动态路阻的计算、道路交通状态的判别等各项功能的实现，都需要实时、准确的短时交通预测数据作为支持。当前，针对交通流量短时预测的方法和模型较多[5]，国内外均有较为丰富的研究[6,7]，分为线性模型、非线性模型、混合模型等类型，主要运用到的方法包括移动平均法、指数平滑法、模糊控制法、基于搜索引擎的方法、灰色理论、卷积神经网络、粒子群预测法、支持向量机等[8~11]。而其中关于时间序列数据的预测模型相对较为成熟，在短时预测方面具有较好的优势。众多学者在基础的时间序列模型上进行了改进，例如唐毅等人[12]利用时间序列模型对高速公路的短时交通流量进行了预测，重点针对参数的动态调整改进了预测模型。姚智胜等人[13]则重点分析了多点时间序列的预测方法。田瑞杰等人[14]则将时间序列数据模型与 BP 神经网络方法进行了结合。整体来看，虽然当前模型及方法均在基本模型上进行了改进，但多为仅支持单一时间尺度的预测，对不同时间尺度下交通信息的特征、差异性、可预测性等考虑不足，导致所构建模型存在较大的误差和不稳定性。为了解决这些问题，本文将重点研究有检测器路段交通流量的多时间尺度预测技术，以提高旅游交通流量预测的准确性。

1　旅游交通流的多时间尺度可预测性分析

在当前的大数据时代，信息预测的方法多种多样，其核

收稿日期：2018-09-02；**修回日期：**2018-11-05　**基金项目：**国家教育部人文社会科学研究项目一般项目（18YJC790084）；中国博士后科学基金面上项目（2018M630281）；河北省社会科学基金资助项目（HB18GL072）；河北省社会科学发展研究课题（201708110301）；燕山大学博士基金资助项目（8190023）

作者简介：李凌雁（1989-），女，河北唐山人，讲师，博士（后），主要研究方向为旅游交通系统分析、交通行为理论；赵欣（1977-），女（通信作者），河北秦皇岛人，教授，博士，主要研究方向为交通规划设计与管理；翁钢民（1963-），男，湖北武人，教授，博导，博士，主要研究方向为区域旅游交通分析与管理.

心均为利用数学方法对样本数据进行处理。要想达到好的预测效果，除选取的预测方法要符合样本数据序列的特性、且样本数据充足之外，还要求样本数据序列本身具备良好的可预测性，即具备一定的统计规律，而非完全随机的数据序列。因此，对动态交通流信息进行多时间尺度预测的前提是确定实时动态交通流参数本身具备可预测性。

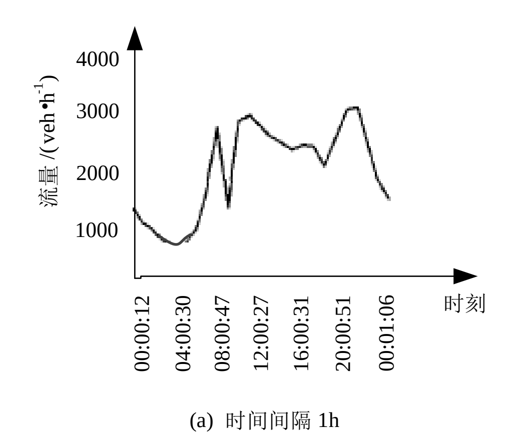
1.1 旅游交通流特性分析

1）旅游交通流的复杂性

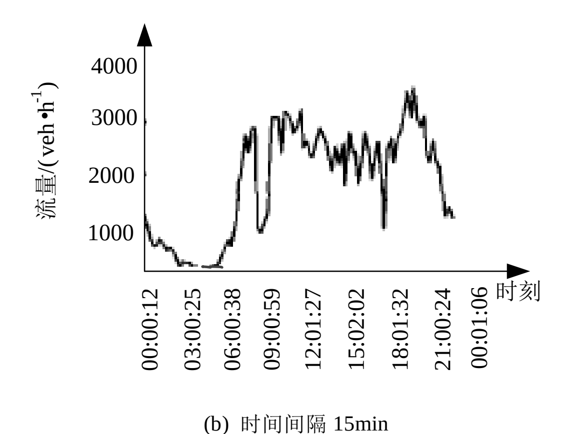
与高速公路相比，城市道路交通流、特别是旅游城市景区交通流具备显著的差异性。主要体现在：高速公路是全封闭、全立交控制出入的，出入口均为匝道控制且相对较少，行车速度稳定可控，而城市道路是由点和线组成的网状结构，单条路段上有多个出入口，具备快速路、主干道、次干道、支路等多种道路类型，且互联互通，行车速度可控性差；高速公路车辆行驶过程中较少更换车道，且方向唯一，而城市道路车流换道、超车行为频繁，且交通流方向多样，选择自由；在不发生交通事故的情况下，高速公路交通流很少产生延误，而城市路网中由于大量信号控制交叉口的存在，交通流时刻处于停停走走状态[15]。此外，城市道路路网较密集，数量多、承担着巨大的交通量，且由于城市布局等原因道路几何条件不统一；存在早晚高峰交通拥挤严重的问题；城市道路中交通流的干扰因素较多，例如行人、自行车过街，公交车停靠等，尤其是景区交通流会显著受到游客量的影响；城市道路还存在左转、直行、右转专用车道。这些特点均导致了旅游道路交通流参数的预测，难度更大、模型选择困难、预测精度较低。

2）多时间尺度的交通流特性分析

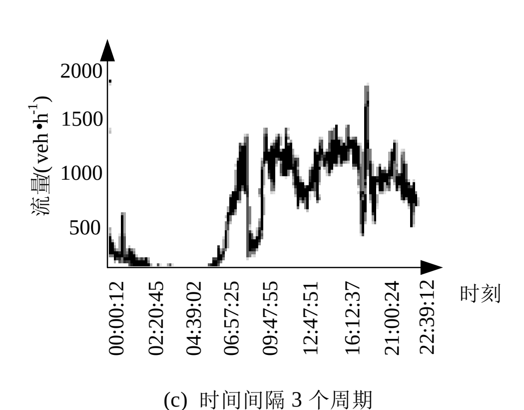
为了探究多时间尺度下旅游交通流的特性，按不同时间尺度对某大型旅游城市滨海大道野生动物园路段的交通参数可预测性进行分析。调查滨海大道主干流向一天 24 h 内的交通流量数据。首先，分别以 1 h、15 min、3 个信号周期、1 个信号周期四种不同的时间尺度为横轴，以小时交通量为纵轴，制作全天交通流量曲线，如图 1 所示。随着四条曲线的时间尺度逐条缩短，交通流量数据的随机性增大，离散性、波动均增强。其次，选取 3 个星期同一天的流量数据进行对比，选择的评价参数有时间序列数据的方差、极值和偏度，计算发现前三种尺度的时间序列数据均具有一定的统计特性，而 1 个信号周期时间尺度的序列数据没有统计特性。可见，应用统计分析方法难以完成交通参数的短时预测。



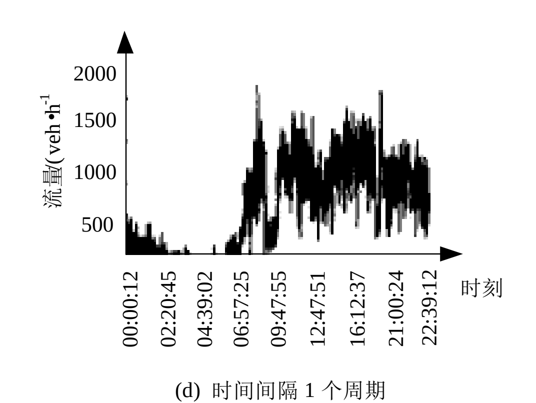
(a) 时间间隔 1h



(b) 时间间隔 15min



(c) 时间间隔 3 个周期



(d) 时间间隔 1 个周期

图 1　多时间尺度的交通数据曲线

Fig. 1　Traffic data curve of multiple time scales

1.2 旅游交通流的多时间尺度可预测性分析

交通流的可预测性可以通过不同观测尺度的流量递归图判断[19]。通过对递归图中与主对角线平行的直线段进行分析，统计线段的长度及密度，能够精准确定旅游交通流量时间序列数据在不同时间尺度下是否具有可预测性。递归图中与主对角线平行的直线段长度越长、密度越大，表明数据越容易收敛、可预测性好、预测精度高；反之，表明数据发散程度高，可预测性不好，预测精度难以保证。

图 2、3 是该旅游城市主干路的流量递归图。拟周期序列(即图 2 的 5 min 时间序列)观测尺度递归图中平行于主对角线的直线段密度较大且相对较长，证明交通流量参数的 5 min 时间尺度可预测性较好，因为 5 min 超过了交叉口的 1 个信号周期，交通流具备规律性。混沌序列(即图 3 的 1 min 时间序列)观测尺度递归图中平行于主对角线的直线段密度较小且相对较短，证明交通流量参数的 1 min 时间尺度可预测性较差，因为 1 min 短于交叉口的 1 个信号周期，交通流离散性明显。

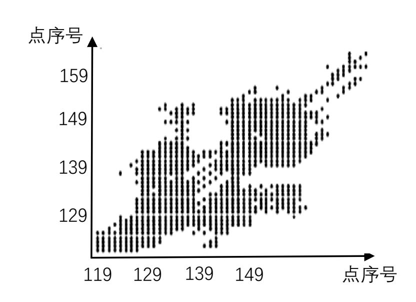


图 2　观测尺度 5 min 时的递归图

Fig. 2　Recurrence plot of 5 min observation scale

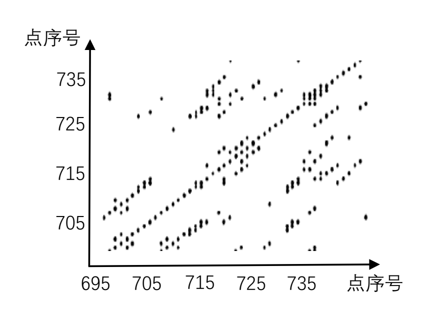


图 3　观测尺度 1 min 时的递归图

Fig. 3　Recurrence plot of 1 min observation scale

2　城市路段旅游交通流量的多时间尺度预测方法

目前，大型旅游城市的主干路段均具备交通信息检测设备，如视频检测器、感应线圈检测器等，交通流量(特别是交叉口的车辆到达率)是可测的[20]。在这种条件下，出现了多种交通流量短时预测方法，根据已知的若干单位时段的交通流量参数，应用加权平均、卡尔曼滤波、神经网络等信息处理手段预测下一个单位时段内交通流量参数的变化趋势。旅游交通管理系统中多模式最优路径的优化、动态路阻的计算、道路交通状态的判别等各项功能对交通预测数据时间尺度上的需求也各不相同(多模式最优路径的优化、动态路阻的计算、道路交通状态的判别对交通预测数据时间尺度上的需求一般分别为 1 min、5 min、15 min)，仅对下一个单位时段内的交通流参数进行预测是不够的，需要同时预测未来若干交通流量参数。

2.1　城市路段旅游交通流量多时间尺度预测模型

1）城市路段旅游交通流量数据的时间序列特性

在城市路网中，交通行为是时空连续行为，动态交通流量参数具备时间序列数据的所有特征，且存在其特殊性。本文主要分析动态交通流量参数的特殊性，以此为基础选择合适的数据挖掘方法。

a）时空关联性。动态交通流量参数同时具备时间关联性与空间关联性。前者体现在同一工作日、同一非工作日的相同地点、相同时段，流量数据是时间相关的，这种特性在各旅游城市采取限号的措施控制交通需求之后表现得更为明显；后者体现在同一交叉口上下游路段上，交通流量数据具备很强的延时空间相关性。

b）交叉口相位协调设计带来的趋势一致性。同一流向的机动车、非机动车流尽量安排在同一相位中，并权衡二者流量确定相位绿灯时长；当各向流量分布不均衡时，可以采用信号相位的早断、迟启等方式来控制非机动车流，且必须对信号相位进行绿灯时间约束，以保证非机动车通行的安全；尽量避免机动车流与非机动车流之间的冲突；同一流向的机动车、非机动车及行人的绿灯信号启亮时间应尽量保持一致。上述措施保证了路网节点上下游同一流向交通流分布的大体均衡。

c）同进口道不同流向的相位设计带来的趋势一致性。确保同一交通流线上相位的连续性；根据交叉口流量分布特性，考虑非机动车流在交叉口的交叉冲突，确定合理的车流组合；根据左转机动车及非机动车的流量比例，协调非机动车左转通行与左转专用相位的设置；重视右转车流在出口处对其他车流及行人的影响，必要时设置右转专用相位。上述措施保证了路网节点上下游不同流向交通流分布的大体均衡[21]。

综上所示，城市旅游交通流量数据是明显的时间序列数据、有一定的规律性和关联性、具备不同时间尺度下的可预测性。本文以相似性搜索为基本思路，以不同时间、空间的交通流量历史数据为基础设计方法，从历史数据库中搜索与未来某时段交通特征相似程度最大的数据序列，结合当前实时采集的数据提出城市路段旅游交通流量多时间尺度预测方法。

2）预测方程的建立

预测方程基本参数定义如下：

a)定义 $Q=\{q_1, q_2, \cdots, q_t\}$为当前日期检测器实时采集的交通流量参数时间序列，其中 $q_1, q_2, \cdots, q_t$ 分别为 1, 2,⋯, t 时段的参数值，每个时段的时间间隔取决于传感器数据采集的时间粒度，也可根据应用功能人为确定，t 代表当前时段。

b)定义 $QB=\{Q_1, Q_2, \cdots, Q_M\}$为历史数据库中交通流量参数时间序列集合，$M$ 代表历史数据库中包含的时间序列数据个数。

c)定义 $Q'=\{q'_1, q'_2, \cdots, q'_t, \cdots\}$为从 QB 中搜索得到与 Q 相似度最大的时间序列数据。

d)定义 $QC=\{q_1, q_2, \cdots, q_t, q'_{t+1}, q'_{t+2},\cdots\}$为 Q 的数据发展规律。

由于交通行为及交通环境是时刻随机变化的，实际的交通系统是一个典型的复杂系统，即使针对同一时空目标，也不可能出现两个完全相同的时间系列数据，也就是说，Q 与 Q'的数据值与数据发展规律是无法完全一致的，需要时刻根据实测数据对 Q'进行修正，以降低预测误差，具体思路如下：

a)应用指数平滑方法预测 $\hat{q}_{t+1}$(t+1 时刻的交通流量参数预测值)，见式（1），其计算依据是 Q 和 q'_{t+1}。

b)同理预测 $\hat{q}_{t+2}$。

c)最终能够得到 $\hat{Q}=\{\hat{q}_{t+1}, \hat{q}_{t+2}, \cdots, \hat{q}_{t+N}\}$，输出预测结果。

$$S_{i+1}=\alpha y_{i+1}+(1-\alpha)S_i$$
$$\hat{q}_{i+1}=S_{i+1}$$
$$\begin{cases}1\le i<t, & y_i=q_i\\ i>t, & y_i=q'_i\end{cases} \tag{1}$$

其中：α 为指数平滑参数，$0<\alpha<1$。

总结上述过程，得到城市路段旅游交通流量多时间尺度预测流程图，如图 4 所示。

3）最大时间尺度的确定方法

理论上，预测步数不能无限制的增加。假设数据采集的时间粒度为 5 min，应用 12：00 的数据预测 12：05、12：10、

12：15 的数据是合理的，随着步数的增加，预测误差随之增大，用来预测时间间隔较大(如 12：50)的数据可能会失去合理性，显然 12：50 的数据应用 12：45 的数据预测会更加合理。所以，对于多时间尺度预测来说，一定存在一个最大的时间尺度，对应最大的预测步数。

假定 N 为与最大时间尺度对应的预测步数，其合理值可由下述方法确定：

a)基于功能需求的定尺度方法。由于旅游交通管理中多模式最优路径的优化、动态路阻的计算、道路交通状态的判别等各项功能的具体实现方法不同，对交通流量预测数据时间尺度上的需求也各不相同，因此，最大时间尺度可根据具体应用功能确定。

b)基于数据计算的变尺度方法。计算当前交通流量时间序列数据与历史时间序列数据的累计相似度，动态调整最大时间尺度。相似度越大，时间尺度可以正向调整，调整量取决于功能能够接受的最大误差。

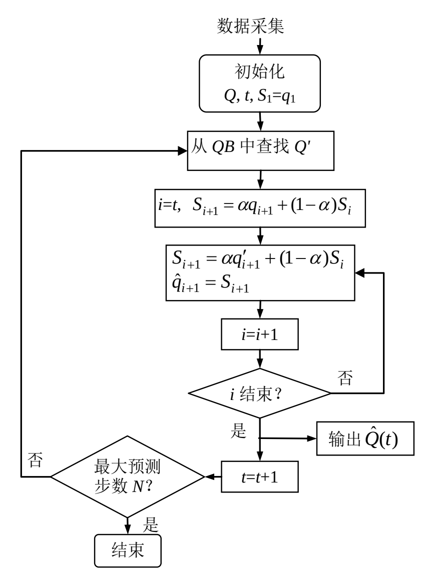


图 4　交通流量多时间尺度预测方法流程

Fig. 4　Process of traffic flow multi time scale prediction

2.2　相似序列的搜索方法

为了保证预测模型的顺利实施，需要设计有效的相似序列搜索方法，快速、准确地找到与预测目标相匹配的数据序列，进而根据上文所构建多时间尺度预测模型进行旅游交通流量的精确预测。设计有效的相似序列搜索方法的关键在于交通流量时间序列数据的表达模式、相似性计算、相似性搜索。

1）交通流量时间序列数据的表达模式

现有文献中的时间序列数据表达模式主要包括符号模式、频域模式、奇异值模式以及分段线性模式[16~18]。通过比较各种表达模式的优缺点，为了保证搜索的实时性，结合交通流量数据的时空属性特点，选取分段线性模式作为交通流量时间序列数据的表达模式。

时间序列分段线性模式表示也可以通过流量递归图直接表达。该模式的核心是如何合理选择直线段段数 *K*，最大程度上保持原有时间序列数据的特性。取值过大会以几何级数增大后续算法的复杂度；取值过小会严重影响相似性搜索的准确性。本文采用“标志点技术”确定 *K* 值。具体步骤如下：

a)确定时间序列数据的一阶标志点。若时间序列 *S* 含有 *n* 个点，假定 *S* 的各个值是时间的函数，则 S 可表示为

$$S_t = y(t)\text{,t=1,2,3} \tag{2}$$

$$\left.\frac{dy(t)}{dt}\right|_{t=t_i} = 0$$

其中：*S* 为时间序列数据，是以时间为变量的函数；S_t 为时间序列的一个数据点，对应时刻 *t*；t_i 为时间序列数据的一阶标志点，定义为$(t_i, y(t_i))$。

b)确定直线段斜率。由步骤 a)得到两个相邻标志点，由两点确定一条直线段，得到斜率：

$$\alpha_i = \frac{y(t_{i+1}) - y(t_i)}{t_{i+1} - t_i} \tag{3}$$

c)误差估计。整个直线段上的累计误差取决于线段内各个数据点与真实值间的相对误差，其估计方法为

$$\beta_j = y(t_{i,j}) - (y(t_i) + \alpha_i \cdot (t_{i,j} - t_i)) \tag{4}$$

$$B_i = \sqrt{\sum_{j=1}^{M} \beta_j^{\,2}} \Big/ M \tag{5}$$

其中：B_i 为直线段累计误差；$\{\beta_j, j=1,2,\ldots,M\}$为线段内各个数据点与真实值间的相对误差集合；*M* 为数据点个数。

2）交通流量时间序列数据的相似性计算

交通流量时间序列数据的相似性可以量化为相似距离，计算方法如下：将某交通流量时间序列数据进行分段线性化表示，将其等价为一组首尾相连的直线段，直线段的交点定义为断点。假设 *S* 为某交通流量时间序列数据的线性分段结果，其包含四个变量、分段数为 *K*。

$$S \equiv \{STL, STR, SYL, SYR\} \tag{6}$$

其中：*STL STR* 为时间标志，即 **S** 中的第 *i* 条直线段的起始时刻；*SYL SYR* 为数值标志，即 **S** 中的第 *i* 段直线起终点对应的交通流量数据。则两个时间序列数据之间的相似距离计算方法为

$$\bar{Y} = \frac{Y - (Y_{MIN} + Y_{MAX}/2)}{(Y_{MAX} - Y_{MIN}/2)} \tag{7}$$

$$D(A,B) = \sum_{i=1}^{K} \left| \frac{(AYR_i - AYL_i)}{(ATR_i - ATL_i)} - \frac{(BYR_i - BYL_i)}{(BTR_i - BTL_i)} \right| \tag{8}$$

其中：$Y_{MIN} = \min(SYL, SYR)$，$Y_{MAX} = \max(SYL, SYR)$。

3）交通流量时间序列数据的相似性搜索

由于电力、通信、气候条件等外界条件影响，旅游交通流量数据的采集往往会存在丢失、失真等现象，导致时间序列历史数据库中的数值无法保证完全连续和完全真实。所以，在依据当前交通流量时间序列数据在历史数据中进行搜索时，搜索的核心目标为寻找与目标数据序列在同一时间段内相似性最大的历史数据序列，对下一时段的交通流量数据的趋势性规律进行多时间尺度预测。搜索过程中需要注意以下几个细节问题：

a)搜索的实时性。旅游交通流量时间序列数据相似性搜

索的目的是从历史数据库中快速、准确找到与当前交通流量时间序列具有最高相似度的时间序列数据。历史数据库中数据量巨大，遍历搜索工作十分复杂，耗时较长，无法保证搜索速度，交通流量数据的实时预测难以实现。因此，可采用图论中分层、分区的思想，将交通流量历史数据库按时间、空间属性不同划分为多个子集。在相似性搜索过程中，首先根据属性确定子集，在子集中遍历搜索，可以大大提升搜索速度。为了进一步提升实时性，可以设置截尾规则，即确定合理的相似距离阈值，当找到某一时间序列数据与目标序列的相似距离小于该阈值时，停止搜索，输出结果；同时确定合理的搜索时间阈值，当搜索用时超过该阈值时，停止搜索，输出已经搜索过与目标序列相似距离最小的交通流量时间序列数据作为结果。

b)实测时间序列数据与历史数据的时间间隔转换。在旅游交通系统中，不同功能的交通信息传感器进行数据采集的时间粒度是不相同的。例如：感应线圈的数据采集间隔一般为 3~5 s，视频检测器的图像识别间隔一般为 5~15 s。而历史数据库从存储空间及数据用途等方面考虑，数据存储时间间隔一般较长，至少大于 10 min。所以，在调用历史数据库进行相似性搜索之前，需要将实时数据转换为与历史存储数据相同的时间间隔。

c)时间序列数据相似性搜索的合理时间间隔。目前，无论针对何种交通参数，其预测一般是短时预测。由于在历史数据库中遍历搜索相似性序列并计算需要一定时间，而传感器的采样时间粒度较短，无法保证对所有样本进行相似性搜索，且没有必要，因为各传感器相邻时段的采集数据变化普遍不大。但如果时间间隔过长，上一时段的相似性搜索结果显然会存在累积误差。所以，为了同时保证预测精度和模型求解实时性，需要合理确定时间序列数据相似性搜索的时间间隔。根据前文分析的旅游交通实际特点，确定该时间间隔为不小于 15 min。

2.3　预测模型的性能评价方法

本文建立的旅游高峰期景区周边路段交通流量参数预测模型是一种微观模型，需要较高的数据预测精度，否则会影响应急交通管理控制策略、方法的实施效果。而精确度主要通过模型预测数据与真实数据之间的误差来反映。绝对误差主要用于测量准确值与预测值之间的差距，相对误差则能够反映模型所得结果的绝对误差与预测值之间的百分比，较好的体现模型的可信程度。在构建旅游交通流预测模型基础上，绝对误差与相对误差的测算最为符合本文研究模型的精确度检验。式(9)为绝对误差的计算方法，式(10)为相对误差的计算方法。

$$S_A(t)=y(t)-\hat{y}(t) \tag{9}$$

$$S_R(t)=\frac{y(t)-\hat{y}(t)}{y(t)} \tag{10}$$

其中：$S_A(t)$为 t 时刻模型的绝对预测误差；$S_R(t)$为 t 时刻模型的相对预测误差；$y(t)$为 t 时刻检测器实际上传的交通流量数据；$\hat{y}(t)$ 为 t 时刻模型输出的交通流量数据预测结果。

在本文建立的预测模型中，某一路段某一特定时刻的交通流量参数预测是多时间尺度的，假设预测步数为 n，模型能够同时输出 n 个预测数据：$\hat{y}_t(t+1)$，$\hat{y}_t(t+2)$，…，$\hat{y}_t(T)$。

为了准确评价多步预测的综合效果，将 n 个预测数据 $\hat{y}_t(t+1)$，$\hat{y}_t(t+2)$，…，$\hat{y}_t(T)$ 的绝对误差与相对误差分别计算平均值作为最终的误差计算公式为

$$MS_A(t)=\frac{1}{n}\sum_{i=t}^{T}{}_t\left(y(i)-\hat{y}_t(i)\right) \tag{11}$$

$$MS_R(t)=\frac{1}{n}\sum_{i=t}^{T}\left(\frac{y(i)-\hat{y}_t(i)}{y(i)}\right) \tag{12}$$

其中：$MS_A(t)$为 t 时刻模型多时间尺度的平均绝对预测误差；$MS_R(t)$为 t 时刻模型多时间尺度的平均相对预测误差；$y(i)$ 为 t 时刻检测器实际上传的交通流量数据。

3　实例分析

本文使用实际数据对城市路段交通流量参数多时间尺度预测方法实施效果进行检验。数据来源于该旅游城市旅游主干道，选取总长 7.66 km 的实验路段，共设置 15 个数据采集点(即通过不间断录像、后期图像处理完成流量数据采集)，数据采集点的平均间距为 0.5 km，后期处理过程中，每个数据采集点以 2 s 为时间粒度统计流量数据。采样时间为 2015 年 6 月 3 日到 2015 年 6 月 8 日。

随机选择编号为 008 的数据采集点为分析目标，由于视频检测器图像处理的技术限制，原始数据序列中存在数据不连续、数据明显异常等现象，对原始数据序列进行了质量评价与控制，具体包括数据修补与数据平滑操作，进行完成了连续时间序列数据的标准化处理工作，结果如图 5 所示。

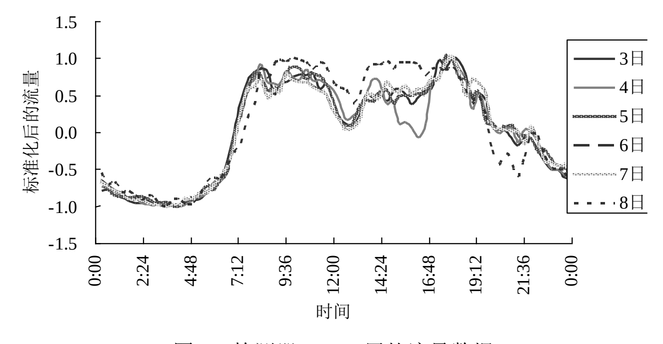


图 5　检测器 008 一周的流量数据

Fig. 5　Flow data in a week of detector 03008

以 008 号数据采集点 2015 年 6 月 3 日的流量时间序列数据为目标，应用 2.2 中构建的相似序列搜索方法从历史数据库中遍历搜索相似序列，应用 2.1 中建立的模型完成预测工作。具体步骤如下：

a)将 2015 年 6 月 3 日的流量原始时间序列数据进行标准化，进而将标准化后的数据进行分段线性化表示，处理结果如图 6 所示。

b)在图 6 中获取 10 个分段点的几何坐标，在历史数据库中提取其余数据采集日内与 10 个断点坐标对应的流量数据，分别计算历史数据与当日数据之间的分段相似度，将各分段点的相似度计算结果逐个累加之后得到累计相似度数据，计算结果如表 1 所示。

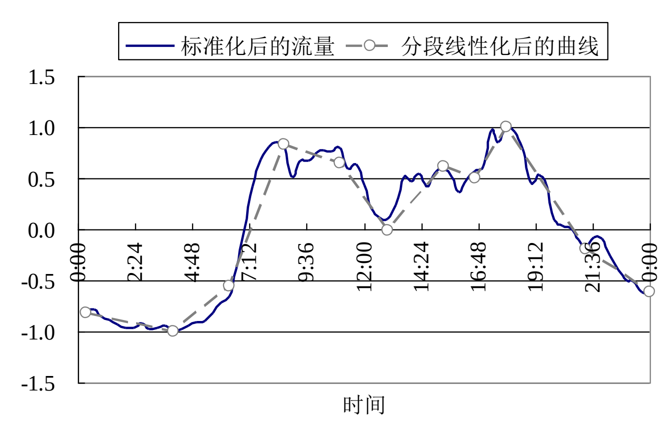


图 6　流量时间序列数据的分段线性化表示

Fig. 6　Piecewise linearization of flow time series data

表 1　交通流量时间序列之间的相似距离

Table 1　Similar distance in traffic flow time series

时间	3 日-4 日		3 日-5 日		3 日-6 日		3 日-7 日		3 日-8 日	
	分段	累计	分段	累计	分段	累计	分段	累计	分段	累计
4:00	0.013	0.013	0.014	0.014	0.016	0.016	0.016	0.016	0.028	0.028
6:20	0.007	0.020	0.008	0.023	0.007	0.023	0.009	0.025	0.008	0.037
8:40	0.042	0.063	0.058	0.081	0.047	0.071	0.059	0.084	0.033	0.069
11:00	0.049	0.112	0.029	0.110	0.057	0.128	0.055	0.139	0.047	0.116
13:00	0.037	0.148	0.017	0.127	0.002	0.130	0.040	0.179	0.042	0.158
15:00	0.114	0.262	0.008	0.135	0.001	0.132	0.003	0.182	0.015	0.173
16:40	0.044	0.307	0.060	0.195	0.014	0.145	0.022	0.203	0.013	0.186
18:00	0.050	0.357	0.042	0.237	0.030	0.176	0.065	0.268	0.095	0.282
21:20	0.031	0.388	0.029	0.266	0.025	0.200	0.006	0.275	0.018	0.300
24:00	0.036	0.424	0.033	0.299	0.022	0.222	0.002	0.277	0.048	0.347

在表 1 中，以 3 日~8 日数据列为例，数据 0.047 代表上午 11 点时刻 3 日断点数据与 8 日断点数据之间的分段相似距离；数据 0.116 代表零点到 11 点时段，3 日与 8 日之间交通流量时间序列数据的分段相似距离加和。数值越小，表示二者的相似程度越高。

c)为了方便观察，应用 Excel 作出样本数据时间序列间分段相似距离与累计相似距离变化曲线图如图 7、8 所示。

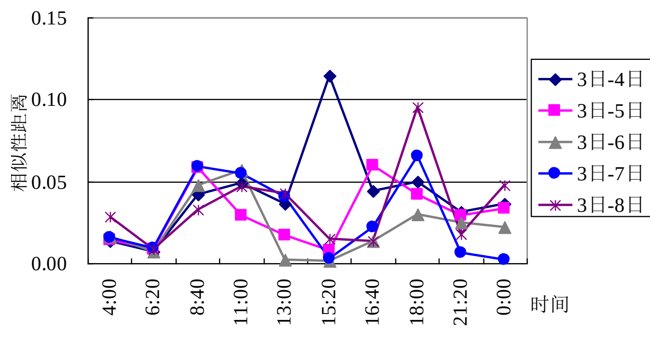


图 7　时间序列间分段相似距离变化曲线

Fig. 7　Change curve of time series piecewise similar distance

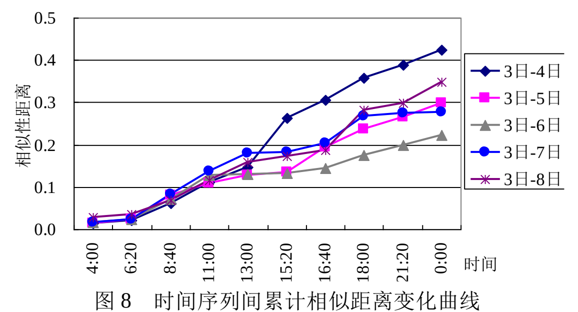


图 8　时间序列间累计相似距离变化曲线

Fig. 8　Change curve of time series cumulative similar distance

显然，与分段相似性距离相比，累计相似距离能够反映两个交通流量时间序列数据之间的整体相似性程度，可将其作为历史数据库搜索的唯一依据。

d)根据图 4 中给出的交通流量多时间尺度预测方法流程对 6 月 3 日 008 号检测点的未来交通流量数据进行多时间尺度预测，选择 15min 时间尺度为例进行误差论证，其预测值与测量值的对比如图 9 所示。

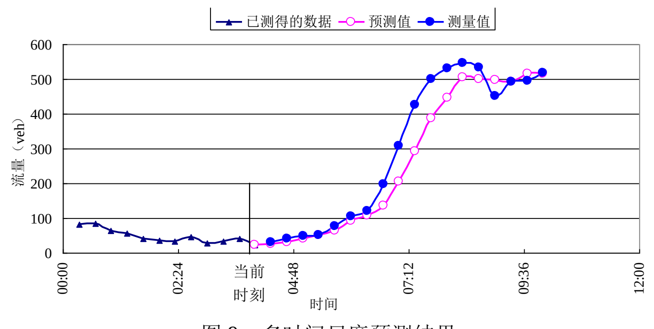


图 9　多时间尺度预测结果

Fig. 9　Prediction of multiple time scales

根据数据计算 15 min 时间尺度预测数据与真实值相比的绝对误差与相对误差，如表 2 所示。由最后一行数据可以看出，城市路段旅游交通流量多时间尺度预测方法在 15 min 时间尺度的预测平均绝对误差为 16.8 辆，平均相对误差为 12.8％，其精度完全满足旅游高峰期交通管理系统的应用需求。

表 2　交通流量 15 min 预测误差结果

Table 2　Prediction error of 15 min traffic flow

时间	误差指标	
	MS_A (veh)	MS_R
4:00	32.364	0.133
6:20	31.393	0.125
8:40	27.474	0.117
11:00	25.179	0.120
13:00	13.589	0.150
15:20	10.128	0.161
16:40	7.835	0.146
18:00	6.757	0.130
21:20	5.595	0.124
平均	16.806	0.128

4　结束语

结合旅游交通需求的具体特点，分析旅游交通预测方法。在确定旅游高峰期交通流量时间序列数据具备可预测性的基础上，提出了基于历史时间序列数据的相似性搜索方法，建立了城市路段旅游交通流量多时间尺度预测模型，并对预测模型的应用效果进行了性能评价。结合实际地区的旅游景区道路实测交通数据对所提出预测模型精度进行了验证。

参考文献：

[1] 张文斌，王博，陈先义．旅游公路交通量预测模型研究[J]．公路工程, 2013, 38(2): 87-90. (Zhang Wenbin, Wang Bo, Chen Xianyi. Study on the traffic forecast model of tourist highway [J]. Highway Engineering, 2013, 38(2): 87-90.)

[2] 黄柯，祝建军，蒲素.我国旅游交通发展现状及研究述评[J]．人文地理, 2007, 22(1): 23-27. (Huang Ke, Zhu Jianjun, Pu Su. Study on the development situation and review of tourism traffic in China [J]. Human Geography, 2007, 22(1): 23-27.)

[3] 余青，韩淼. 城郊型风景名胜区旅游交通治堵 PMSI 模式及实践应用——以雁荡山—楠溪江国家风景名胜区为例[J]. 地理研究，2016, 35(8):1589-1598. (Yu Qing, Han Miao. Construction and practical application of PMSI model for tourist traffic congestion in suburban scenic area: A case study of Yandangshan-Nanxijiang National Scenic Area [J]. Geographical Research, 2016, 35(8): 1589-1598.)

[4] Sung S P. A linear expenditure model for tourism demand [J]. Annals of Tourism Research, 2011(8): 443-454.

[5] Sevgin A. A compact econometric model of tourist demand [J]. Tourism Management, 2008, 19(1): 99-102.

[6] Richard J, Anne P. Research on the method used in predicting the traffic volume on non-detector intersections [C]//Proc of the 11th World Congress on ITS.2014:122-129.

[7] Yang D. Combined traffic signal control and route guidance: multiple user class traffic assignment model versus discrete choice model [J]. Tourism Management, 2005, 26(3): 359-365.

[8] 傅成红，张阳. 基于参数优化的 SVR 城市群交通需求预测方法[J]. 系统工程，2016(2): 114-120. (Fu Chenghong, Zhang Yang. An improved SVR algorithm based on the nuclear parameter optimization for traffic demand forecasting of urban agglomeration [J]. Systems Engineering, 2016(2): 114-120.)

[9] 陈曦，彭蕾，李炜. 基于计算实验的公共交通需求预测方法 [J]. 自动化学报，2017, 43(1): 60-71. (Chen Xi, Peng Lei, Li Wei. A computational experiment approach to public traffic demand forecast [J]. Acta Automatica Sinica, 2017, 43(1): 60-71.)

[10] 张兴平，杨建军，毛必林. 杭州市区旅游交通网络空间分析及其对策 [J]. 浙江大学学报:理学版，2000, 27(4): 438-443. (Zhang Xingping, Yang Jianjun, Mao Bilin. Spatial analysis and countermeasures of tourism transportation network of Hangzhou urban area [J]. Journal of Zhejiang University:Sciences Edition, 2000, 27(4): 438-443)

[11] 梅朵，郑黎黎，冷强奎，等.基于时空 GPSO-SVM 的短时交通流预测[J]. 交通信息与安全, 2017, 35(2): 68-74. (Mei Duo, Zheng Lili, Leng Qiangkui, *et al*. A prediction model for short-term traffic flow based on space-time GPSO-SVM [J]. Journal of Transport Information and Safety, 2017, 35(2): 68-74.)

[12] 唐毅，刘卫宁，孙棣华，等. 改进时间序列模型在高速公路短时交通流量预测中的应用 [J]. 计算机应用研究, 2015, 32(1):146-149. (Tang Yi, Liu Weining, Sun Dihua, *et al*. Application of improved time series model in forecasting of short-term traffic flow for freeway [J]. Application Research of Computers, 2015, 32(1): 146-149.)

[13] 姚智胜，邵春福. 基于状态空间模型的道路交通状态多点时间序列预测 [J]. 中国公路学报, 2007, 20(4): 113-117. (Yao Zhisheng, Shao Chunfu. Road traffic state multi-spot time series forecasting based on state space model [J]. China Journal of Highway and Transport, 2007, 20(4): 113-117.)

[14] 田瑞杰，张维石，翟华伟. 基于时间序列与 BP-ANN 的短时交通流速度预测模型研究[J/OL]. 计算机应用研究, 2019, 36(11). http://www.arocmag.com/article/02-2019-11-033.html. (Tian Ruijie, Zhang Weishi, Zhai Huawei. Short-term traffic flow velocity prediction model based on time series and BP-ANN [J/OL]. Application Research of Computers, 2019, 36 (11). http://www.arocmag.com/article/02-2019-11-033.html .)

[15] 陈志杰，毛保华，柏赟，等. 基于多时间尺度的城市轨道交通短时 OD 估计[J]. 交通运输系统工程与信息，2017, 17(5): 166-172. (Chen Zhijie, Mao Baohua, Bai Yun, *et al.* Short-term origin-destination estimation for urban rail transit based on multiple temporal scales [J]. Journal of Transportation Systems Engineering and Information Technology, 2017, 17(5): 166-172.)

[16] 罗向龙，焦琴琴，牛力瑶，等. 基于深度学习的短时交通流预测 [J]. 计算机应用研究，2017, 34(1): 91-93. (Luo Xianglong, Jiao Qinqin, Niu Liyao, Sun Zhuangwen. Short-term traffic flow prediction based on deep learning [J]. Application Research of Computers, 2017, 34(1): 91-93.)

[17] Linet O, Ediz E, Beste K. Emergency Logistics Planning in Natural Disasters [J]. Annals of Operations Research, 2004, 129: 217-245.

[18] 杨刚，王乐，戴丽珍，等. 自适应权重粒子群优化 LS-SVM 的交通流预测 [J]. 控制工程, 2017, 24(9): 1838-1843. (Yang Gang, Wang Le, Dai Lizhen, *et al.* Traffic Flow Prediction Based on Adaptive Particle Swarm Optimization of Least Square Support Vector Machine [J]. Control Engineering of China, 2017, 24(9): 1838-1843.)

[19] 李松，刘力军，翟曼. 改进粒子群算法优化 BP 神经网络的短时交通流预测 [J]. 系统工程理论与实践, 2012, 32(9): 2045-2049. (Li Song, Liu Lijun, Zhai Man. Prediction for short-term traffic flow based on modified PSO optimized BP neural network [J]. System Engineering-Theory & Practice, 2012, 32(9): 2045-2049.)

[20] 贾宁，马寿峰，钟石泉. 基于遗传算法优化和 KD 树的交通流非参数回归预测方法 [J]. 控制与决策, 2012, 27(7): 991-996. (Jia Ning, Ma Shoufeng, Zhong Shiquan. Non-parameter-regression traffic flow forecast method based on KD-tree and genetic optimization [J]. Control and Decision, 2012, 27(7): 991-996.)

[21] 韩志聪，樊彦国，吴会胜，等. 基于 GA-SVR 模型的短期交通流量预测方法研究 [J]. 公路交通科技, 2017, 34(1): 130-136. (Han Zhicong, Fan Yanguo, Wu Huisheng, *et al.* Study on methods of short-time traffic volume forecasting based on GA-SVR model [J]. Journal of Highway and Transportation Research and Development, 2017, 34(1): 130-136.)